Journal of Multivariate Analysis 110 (2012) 91–105

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Comparison of estimators for pair-copula constructions

Ingrid Hobæk Haff

Norwegian Computing Center, PB 114 Blindern, NO-0373 Oslo, Norway

ARTICLE INFO

Article history:
Available online 23 August 2011

AMS subject classifications:
62L12
62H12
62F10
62G30

Keywords:
Semiparametric estimation
Copulae
Hierarchical structures
Empirical distribution functions
Simulation study

ABSTRACT

We compare two of the most used estimators for the parameters of a pair-copula construction (PCC), namely the semiparametric (SP) and the stepwise semiparametric (SSP) estimators. By construction, the computational speed of the SSP estimator is considerably higher, at the expense of its asymptotic efficiency. Based on an extensive simulation study, we find that the performance of the SSP estimator is overall satisfactory compared to its contender. SSP loses some efficiency with respect to SP with increasing dependence, especially in the top levels of the PCC. On the other hand, the SSP estimator may suffer less under reduced sample sizes and misspecification of the model. Finally, it is the only real alternative for large-dimensional problems. Though it struggles with the top level parameters, the lower order dependences of the resulting estimated PCC mimic the true distribution well. All in all, this study supports the use of SSP in most applications.

1. Introduction

Multivariate modeling is a key field in statistics. In particular, the interest in copulae has boomed since the late 1990's [11], probably due to the books by Joe [19] and Nelsen [26], as well as the introduction of copulae to financial applications by Frees and Valdez [9] and Embrechts et al. [8]. It has entailed the development of a number of hierarchical, copula-based structures, among those the pair-copula construction (PCC), originally proposed by Joe [18]. This structure has later been explored by Bedford and Cooke [2,3] and Kurowicka and Cook [25] from a probabilistic point of view, and by Aas et al. [1] for inference.

PCCs are treelike constructions, having pair-copulae as building-blocks. Delightfully simple as these structures are, they no doubt owe their popularity to high flexibility and the ability to model a wide scope of dependences [21,17]. Accordingly, several estimators for PCC parameters have been proposed, e.g. [22,1,7,24] and more recently, [15]. The aim of this work is to compare alternative estimation procedures. A PCC is in fact a multivariate copula. All standard copula parameter estimators are therefore applicable also to PCCs. We are, however, not that interested in the rather model specific method of moments type estimators, typically, an inversion of Kendall's τ coefficients [6,27,10,14].

Among the more general approaches, the most relevant are semiparametric (SP) estimation, inference function for margins estimation and maximum likelihood, of which the latter two depend on the specified margins. Our focus is on the dependence parameters. Moreover, the effect of margins on the estimation of copula parameters has already been extensively studied [20,23]. Hence, out of the above three, we will only consider the semiparametric estimator, that will serve as a benchmark. Along with the contending stepwise semiparametric (SSP) estimator, it is one of the most commonly used estimators for PCCs.

As one would expect, the malleability of pair-copula constructions does not come without a price. Even low-dimensional structures have many parameters, and the number grows quickly with the number of variables. SP estimation consists in estimating all parameters simultaneously, which requires numerical optimization. Hence, it is very likely that it will be

E-mail addresses: ingrid@nr.no, Ingrid.Haff@nr.no.

0047-259X/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2011.08.013

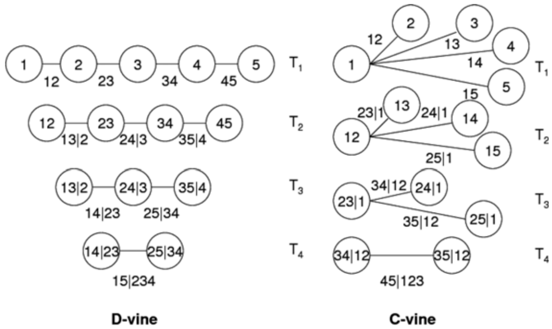


Fig. 1. Five-dimensional D-vine (to the left) and C-vine (to the right).

numerically challenging and time consuming with increasing dimension. The SSP estimator, which is designed for PCCs, performs the estimation in several steps, which increases the computational speed considerably, but reduces its asymptotic efficiency [15].

There exist expressions for the asymptotic covariance matrices of these estimators (see for instance [12,15]). However, they involve multiple integrals, which, in practice, are incalculable. Therefore, and in order to investigate finite sample behavior, we base the comparison on an extensive simulation study.

The paper is organized as follows. Section 2 presents the model, i.e. PCCs, whereas the two estimators are introduced in Section 3. The results of the simulation study are exhibited in Section 4. Finally, we summarize and discuss the results in Section 5.

2. Model

Consider a d-variate stochastic vector $\boldsymbol{X} = (X_1, \ldots, X_d)^\top$ from an absolutely continuous distribution $F_{1\cdots d}$ with strictly increasing margins $F_1, \ldots, F_d$. Using Sklar's theorem [29], as well as the chain rule, the probability density function (pdf) of $\boldsymbol{X}$ may be expressed as

$$f_{1\cdots d}(x_1, \ldots, x_d) = \cdot c_{1\cdots d}(F_1(x_1), \ldots, F_d(x_d)) \prod_{\ell=1}^{d} f_\ell(x_\ell), \tag{1}$$

where $f_\ell,\ \ell = 1, \ldots, d$, are the corresponding marginal pdfs and $c_{1\cdots d}$ the copula density. Likewise, it can be factorized as

$$f_{1\cdots d}(x_1, \ldots, x_d) = f_1(x_1) f_{2|1}(x_2|x_1) \cdots f_{d|1\cdots d-1}(x_d|x_1, \ldots, x_{d-1}). \tag{2}$$

The related pair-copula construction (PCC) results from expressing the factors on the right-hand side of (2) in terms of pair-copula densities and marginal pdfs, through the repeated use of (1). Applying (1) in two dimensions, the second factor $f_{2|1}(x_2|x_1)$ is given by

$$f_{2|1}(x_2|x_1) = f_2(x_2) c_{12}(F_1(x_1), F_2(x_2)).$$

Correspondingly,

$$\begin{aligned} f_{3|12}(x_3|x_1, x_2) &= f_{3|2}(x_3|x_2) c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \\ &= f_3(x_3) c_{23}(F_2(x_2), F_3(x_3)) c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)), \end{aligned}$$

where $c_{13|2}$ is the copula density for the distribution of the pair (X_1, X_3) conditioning on X_2. As one continues with the remaining factors of (2), one finally obtains a product of marginal pdfs and pair-copula densities, i.e. a PCC.

All components of the structure, both marginal and pair-copula densities, can be chosen freely, i.e. from different families. The resulting distribution is guaranteed to be valid. Despite its simple building blocks, the pair-copula construction is therefore an exceptionally flexible model, able to portray a wide range of dependence structures [21].

There are numerous ways of factorizing $f_{1\cdots d}$ and of substituting the factors with pair-copula densities, each resulting in a valid PCC. A large subset of these belongs to the family of regular vines, introduced by Bedford and Cooke [2,3], which again comprises canonical (C) and drawable (D) vines. Five-dimensional examples of the latter two are shown in Fig. 1. For a more thorough introduction to vines and pair-copula constructions, see for instance [1].

For simplicity, we will hereafter restrict our attention to D-vines. The corresponding pdf is given by

$$f_{1\cdots d}(x_1, \ldots, x_d) = \prod_{\ell=1}^{d} f_\ell(x_\ell) \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,i+j|v_{ij}} \left(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}), F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}}) \right), \tag{3}$$

where v_{ij} denotes the index set $\{i+1, \ldots, i+j-1\}$. In the double product over the pair-copula densities, j runs over the levels of the structure, for instance T_1–T_5 in Fig. 1, and i over the copulae at each level. Note that excepting the ground level, the arguments of the pair-copulae are conditional distributions, whose number of conditioning variables increases by one with each level. For inference to be possible in practice, one has to make the assumption that these so-called conditional pair-copulae depend on the conditioning variables only through their arguments, thus obtaining a *simplified* PCC. Take for example $c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2))$, which is given by

$$c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) = \frac{f_{13|2}(x_1, x_3|x_2)}{f_{1|2}(x_1|x_2)f_{3|2}(x_3|x_2)}.$$

If $C_{13|2}$ is a Gaussian copula with parameter ρ, one must assume that ρ is constant over all values of x_2. This is generally not the case, but Hobæk Haff et al. [17] showed that this is not a very restricting assumption, and that a general PCC may be very well approximated by a simplified one.

Assume now that we have n independent observations $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ from the model (3), and that we wish to estimate the corresponding parameters. The log-likelihood function is given by

$$\begin{aligned} l(\boldsymbol{\alpha}, \boldsymbol{\theta}; \boldsymbol{x}_1, \dots, \boldsymbol{x}_n) &= \sum_{k=1}^{n}\sum_{\ell=1}^{d} \log f_\ell(x_{\ell k}; \boldsymbol{\alpha}) \\ &\quad + \sum_{k=1}^{n}\sum_{j=1}^{d-1}\sum_{i=1}^{d-j} \log c_{i,i+j|v_{ij}}\left(F_{i|v_{ij}}(x_{ik}|\boldsymbol{x}_{v_{ij},k}; \boldsymbol{\alpha}, \boldsymbol{\theta}), F_{i+j|v_{ij}}(x_{i+j,k}|\boldsymbol{x}_{v_{ij},k}; \boldsymbol{\alpha}, \boldsymbol{\theta}); \boldsymbol{\theta}\right), \end{aligned} \tag{4}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ are the marginal and dependence parameters, respectively. The computation of (4) requires calculation of the conditional distributions that are arguments of the pair-copulae. Using the simplifying assumption, one may express them as functions of two other conditionals with one conditioning variable less. As shown by Joe [19]

$$F_{i|v\cup\{j\}}(x_i|\boldsymbol{x}_{v\cup\{j\}}) = \left.\frac{\partial C_{ij|v}(u_i, u_j)}{\partial u_j}\right|_{u_i=F_{i|v}(x_i|\boldsymbol{x}_v), u_j=F_{j|v}(x_j|\boldsymbol{x}_v)}, \tag{5}$$

where i, j are distinct indices, and v is a non-empty set of indices, that contains neither i nor j. Likewise, $F_{i|v}$ and $F_{j|v}$ can be expressed as bivariate functions of conditional distributions with a conditioning set reduced by one, and so on. Hence, all the necessary conditional distributions in (4) are nested functions of the margins $F_1, \dots, F_d$. The log-likelihood function can therefore be written as

$$l(\boldsymbol{\alpha}, \boldsymbol{\theta}; \boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = l_M(\boldsymbol{\alpha}; \boldsymbol{x}_1, \dots, \boldsymbol{x}_n) + l_C(\boldsymbol{\theta}; \boldsymbol{u}_1(\boldsymbol{\alpha}), \dots, \boldsymbol{u}_n(\boldsymbol{\alpha})), \tag{6}$$

where $\boldsymbol{u}_k(\boldsymbol{\alpha}) = (F_1(x_{1k}; \boldsymbol{\alpha}), \dots, F_d(x_{dk}; \boldsymbol{\alpha}))$.

3. Estimators

The estimators we wish to compare are two of the most commonly used for pair-copula constructions, namely the semiparametric and the stepwise semiparametric estimators.

3.1. The semiparametric estimator

The semiparametric (SP) estimator was introduced by Genest et al. [12], and for censored data by Shih and Louis [28]. Possible extensions were later considered by Tsukahara [30]. Dependence on the chosen margins is removed by replacing the parametric cdfs $F_i(\cdot, \boldsymbol{\alpha})$ in (6) with the empirical ones $F_{in}(\cdot)$. If one is interested in estimating measures, such as Kendall's τ, Spearman's ρ or tail dependence coefficients, that are functions only of the dependence parameters, the SP estimator is a very natural choice. Not only is it more robust to misspecified margins, but specification thereof is completely avoided.

Define the pseudo observations

$$u_{ikn} = F_{in}(x_{ik}) = \frac{1}{n+1}\sum_{j=1}^{n} I(x_{ij} \le x_{ik}), \quad i = 1, \dots, d, k = 1, \dots, n,$$

where $I(\cdot)$ is the indicator function, and the pseudo log-likelihood function as

$$l_{C,P}(\boldsymbol{\theta}; \boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = l_C(\boldsymbol{\theta}; \boldsymbol{u}_{1n}, \dots, \boldsymbol{u}_{nn}), \tag{7}$$

with l_C as defined in (6) and $\boldsymbol{u}_{kn} = (u_{1kn}, \dots, u_{dkn})$. This is the sum over all log-copula densities, plugging in the pseudo observations. The SP estimator is now simply the maximizer of $l_{C,P}$ with respect to $\boldsymbol{\theta}$, i.e.

$$\hat{\boldsymbol{\theta}}^{SP} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\{l_{C,P}(\boldsymbol{\theta}; \boldsymbol{x}_1, \dots, \boldsymbol{x}_n)\}.$$

This estimation procedure is thoroughly described for D-vines in [1], whereas the large sample properties of the estimator for PCCs are described for instance in [15].

Although the pseudo log-likelihood function can be expressed directly as a function of the empirical margins, it is, in practice, computed iteratively, level by level. At a given level, the necessary pair-copula arguments are computed by applying the functions (5) to the arguments from the preceding level. Thus, they are successive transformations of the pseudo observations. More specifically, the arguments of level ℓ are the result of $\ell - 1$ such transformations.

Previous studies have shown that the SP estimator performs well for copulae compared to alternative estimators, such as maximum likelihood, inference function for margins and the minimum distance estimator [23,30]. However, these studies have focused on bivariate examples. As mentioned earlier, the number of parameters of a pair-copula construction grows quickly with the number of variables. For instance, a five-dimensional D-vine consisting of t-copulae has 20 parameters (not counting the margins). With an additional dimension, the number of parameters increases to 30. Thus, for a medium to high number of variables, the optimization of the pseudo log-likelihood function becomes numerically demanding and time consuming.

3.2. The stepwise semiparametric estimator

By performing the estimation in several steps, one can speed up the procedure considerably. This is the main idea of the stepwise semiparametric (SSP) estimator, which is designed for PCCs. It was suggested in [1], and more formally introduced by Hobæk Haff [15]. The latter also explores its large sample characteristics.

The SSP estimator is very similar to the SP one. The difference is that the PCC parameters are estimated level by level, plugging in parameters from previous levels at each step. Let

$$l_{C,\ell} = \sum_{k=1}^{n} \sum_{j=1}^{\ell} \sum_{i=1}^{d-j} \log c_{i,i+j|v_{ij}}.$$

This is the sum of all log-copula densities up to and including level ℓ, over all observations. Note that $l_{C,d-1} = l_C$ for the top level $\ell = d - 1$. Now, let $\boldsymbol{\theta}_\ell$ denote the parameters of all pair-copulae at level ℓ. Then, $l_{C,\ell}$ is a function of the parameters $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_\ell$ from level 1 up to ℓ, but not of $\boldsymbol{\theta}_{\ell+1}, \ldots, \boldsymbol{\theta}_{d-1}$ from the ensuing levels. Analogously to (7), define the ℓ-level pseudo log-likelihood function $l_{C,P,\ell}$ as

$$l_{C,P,\ell}(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_\ell; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = l_{C,\ell}(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_\ell; \boldsymbol{u}_{1n}, \ldots, \boldsymbol{u}_{nn}), \tag{8}$$

i.e. by substituting the parametric margins for the empirical ones. At the top level, (8) is simply the pseudo log-likelihood function from (7), when seen as a function of $\boldsymbol{\theta}$. Nonetheless, the SSP top level estimates are different from the SP ones. The estimation procedure is as follows:

- maximize $l_{C,P,1}(\boldsymbol{\theta}_1; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ over $\boldsymbol{\theta}_1$ to obtain $\hat{\boldsymbol{\theta}}_1^{SSP}$.
- for level $\ell = 2, \ldots, d-1$ maximize $l_{C,P,\ell}(\hat{\boldsymbol{\theta}}_1^{SSP}, \ldots, \hat{\boldsymbol{\theta}}_{\ell-1}^{SSP}, \boldsymbol{\theta}_\ell; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ over $\boldsymbol{\theta}_\ell$ to obtain $\hat{\boldsymbol{\theta}}_\ell^{SSP}$.

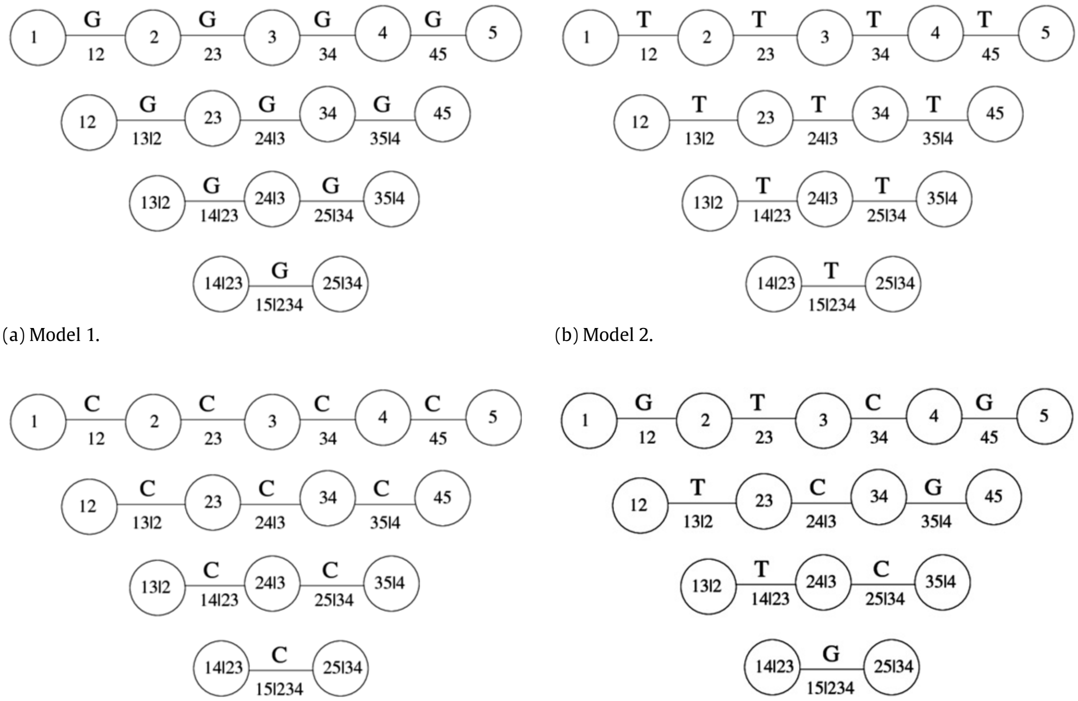
For a detailed estimation algorithm and description of how to compute $l_{C,P,\ell}$, $\ell = 1, \ldots, d-1$, see [15]. Note that when none of the copulae constituting the structure share parameters, as in the models we use for the comparison of estimators (Section 4), the optimization is done for each pair-copula, individually. Also, the necessary pair-copula arguments at a given level are computed by iterative transformations of the pseudo observations, just as for the SP estimator (see Section 3.1). The difference is that the parameters plugged into the transformations are final estimates, while they are part of the simultaneous optimization in SP estimation.

The SSP estimator does not take into account information the next levels might have on the parameters. Obviously, it is asymptotically less efficient than the SP estimator. Nonetheless, it is useful to use SSP in any case, to get starting values for the SP estimation. The question is, how much precision does one really gain by a subsequent SP estimation? Is it worth the extra time spent?

4. Comparison

We want to study how the SP and SSP estimators perform for pair-copula constructions, both in terms of computing time and finite sample bias and variance. As mentioned earlier, we concentrate on the subgroup of D-vines. More specifically, we base all experiments, but one, on five-dimensional D-vines (as depicted in Fig. 1), varying the copula types. As explained earlier, the choice of margins does not affect the dependence parameter estimates since both estimators use the data transformed with their empirical margins. Therefore, we let the margins be uniform $U[0, 1]$ in all the experiments, without loss of generality. Note that we still assume the margins to be unknown in the estimation, and therefore plug the pseudo observations into (7) and (8).

The object of the study is to explore how the estimators' performance is affected by the type and degree of dependence, the number of observations n and the correctness of the model. We also include one large-dimensional example.



(a) Model 1. (b) Model 2. (c) Model 3. (d) Model 4.

Fig. 2. Models used for comparison. The first three, (a)–(c), consist of only one of the three types of copulae, Gaussian (G), Student's t (T) and Clayton (C), respectively, while the last, (d), is a mix of the three.

4.1. Type of dependence

By type, we mean presence, or not, of tail dependence and dependence asymmetry. We account for this by considering three copula families, namely

- the Gaussian (no tail dependence),
- the Student's t (tail dependence)
- and the Clayton copula (lower, but not upper tail dependence),

combined in four different models. The first three of these consist of only one of the copula types, whereas the last is a mix of all three types (see Fig. 2).

Moreover, we have varied the degree of dependence by means of different parameter values. More specifically, we let the correlation parameter ρ of the Gaussian and Student's t copulae take the values $\{0.2, 0.5, 0.8\}$, corresponding to low, medium and high dependence, respectively. To facilitate comparison, we chose to consider the values $\{0.294, 1, 2.88\}$ for the parameter α of the Clayton copulae, that give the same Kendall's τ coefficients as the ρ values above. Furthermore, we fixed the number of degrees of freedom of the Student's t copulae to $\nu = 6$, which ensures a rather strong tail dependence.

In all experiments, we have generated $n = 5000$ independent samples from the model in question, which is a relatively large sample size compared to most applications, estimated the parameters using SP and SSP, and repeated this $N = 1000$ times. The measures we have compared are

- the finite sample bias $Bias(\hat{\theta}_i, \theta_i) = \frac{1}{N}\sum_{k=1}^{N} |\hat{\theta}_{ik} - \theta_i|$,
- the mean squared error $MSE(\hat{\theta}_i, \theta_i) = \frac{1}{N}\sum_{k=1}^{N} (\hat{\theta}_{ik} - \theta_i)^2$
- and the CPU time (in seconds),

the former two being computed for each parameter θ_i, $i = 1, \ldots, |\boldsymbol{\theta}|$. All computations were programmed in *R*, version 2.11.1, using the function `optim()` for the optimizations. Further, they were run on a computer with 72 GB RAM and an E5530 @ 2.40 GHz processor with 8 CPU kernels and hyperthreading, that allows for 16 simultaneous threads.

The results are summarized in Table 1 for Model 2(a)–(c), and in Table 2 for Model (d). In the former table, the results are averaged over each of the four levels of the structure, while they are presented per copula in the latter. Starting with the models consisting of copulae of the same type (Table 1), the bias and MSE of two estimators decrease with the degree of dependence. The reason is probably that the log-likelihood function steepens at stronger dependence. Further, the two measures appear to be rather constant over the four levels at low and medium dependence, while they increase with the level at high dependence. As explained in Section 3, the arguments of pair-copulae above the ground level are obtained through

Table 1
Finite sample bias and MSE of $\boldsymbol{\theta}$ based on simulations of $n = 5000$ observations from Model 2(a)–(c) (see Fig. 2), i.e. consisting of only one copula type, with low ($\rho = 0.2$, $\alpha = 0.294$), medium ($\rho = 0.5$, $\alpha = 1$) and high ($\rho = 0.8$, $\alpha = 2.88$) dependence, as well as CPU time in seconds.

			SP					
Degree of dependence			Low		Medium		High	
Model	Par.	Level	Bias	MSE	Bias	MSE	Bias	MSE
Gaussian	ρ	1	$1.09 \cdot 10^{-2}$	$1.85 \cdot 10^{-4}$	$8.49 \cdot 10^{-3}$	$1.12 \cdot 10^{-4}$	$3.96 \cdot 10^{-3}$	$2.44 \cdot 10^{-5}$
		2	$1.09 \cdot 10^{-2}$	$1.87 \cdot 10^{-4}$	$8.31 \cdot 10^{-3}$	$1.08 \cdot 10^{-4}$	$4.16 \cdot 10^{-3}$	$2.75 \cdot 10^{-5}$
		3	$1.09 \cdot 10^{-2}$	$1.86 \cdot 10^{-4}$	$8.82 \cdot 10^{-3}$	$1.20 \cdot 10^{-4}$	$4.26 \cdot 10^{-3}$	$2.88 \cdot 10^{-5}$
		4	$1.09 \cdot 10^{-2}$	$1.85 \cdot 10^{-4}$	$8.41 \cdot 10^{-3}$	$1.13 \cdot 10^{-4}$	$9.48 \cdot 10^{-3}$	$1.17 \cdot 10^{-4}$
		CPU	$1.79 \cdot 10^{-2}$		$5.33 \cdot 10^{-2}$		$4.05 \cdot 10^{-1}$	
Student's t	ρ	1	$1.18 \cdot 10^{-2}$	$2.18 \cdot 10^{-4}$	$9.31 \cdot 10^{-3}$	$1.35 \cdot 10^{-4}$	$4.64 \cdot 10^{-3}$	$3.42 \cdot 10^{-5}$
		2	$1.19 \cdot 10^{-2}$	$2.19 \cdot 10^{-4}$	$9.30 \cdot 10^{-3}$	$1.36 \cdot 10^{-4}$	$4.71 \cdot 10^{-3}$	$3.43 \cdot 10^{-5}$
		3	$1.19 \cdot 10^{-2}$	$2.18 \cdot 10^{-4}$	$9.26 \cdot 10^{-3}$	$1.35 \cdot 10^{-4}$	$4.74 \cdot 10^{-3}$	$3.57 \cdot 10^{-5}$
		4	$1.20 \cdot 10^{-2}$	$2.27 \cdot 10^{-4}$	$9.75 \cdot 10^{-3}$	$1.49 \cdot 10^{-4}$	$1.09 \cdot 10^{-2}$	$1.53 \cdot 10^{-4}$
	ν	1	$4.88 \cdot 10^{-1}$	$3.79 \cdot 10^{-1}$	$3.96 \cdot 10^{-1}$	$2.54 \cdot 10^{-1}$	$3.86 \cdot 10^{-1}$	$2.41 \cdot 10^{-1}$
		2	$4.81 \cdot 10^{-1}$	$3.76 \cdot 10^{-1}$	$4.29 \cdot 10^{-1}$	$3.07 \cdot 10^{-1}$	$3.10 \cdot 10^{-1}$	$1.56 \cdot 10^{-1}$
		3	$4.96 \cdot 10^{-1}$	$4.04 \cdot 10^{-1}$	$4.58 \cdot 10^{-1}$	$3.51 \cdot 10^{-1}$	$3.08 \cdot 10^{-1}$	$1.52 \cdot 10^{-1}$
		4	$5.09 \cdot 10^{-1}$	$4.28 \cdot 10^{-1}$	$5.17 \cdot 10^{-1}$	$4.50 \cdot 10^{-1}$	$3.01 \cdot 10^{-1}$	$1.45 \cdot 10^{-1}$
		CPU	1.66		1.97		1.08	
Clayton	α	1	$1.66 \cdot 10^{-2}$	$4.28 \cdot 10^{-4}$	$2.65 \cdot 10^{-2}$	$1.09 \cdot 10^{-3}$	$5.91 \cdot 10^{-2}$	$5.41 \cdot 10^{-3}$
		2	$1.62 \cdot 10^{-2}$	$4.21 \cdot 10^{-4}$	$2.37 \cdot 10^{-2}$	$8.94 \cdot 10^{-4}$	$5.68 \cdot 10^{-2}$	$4.85 \cdot 10^{-3}$
		3	$1.67 \cdot 10^{-2}$	$4.31 \cdot 10^{-4}$	$2.63 \cdot 10^{-2}$	$1.10 \cdot 10^{-3}$	$1.24 \cdot 10^{-1}$	$1.95 \cdot 10^{-2}$
		4	$1.67 \cdot 10^{-2}$	$4.46 \cdot 10^{-4}$	$2.78 \cdot 10^{-2}$	$1.24 \cdot 10^{-3}$	$6.12 \cdot 10^{-1}$	$3.83 \cdot 10^{-1}$
		CPU	$1.88 \cdot 10^{-2}$		$6.49 \cdot 10^{-2}$		$6.07 \cdot 10^{-2}$	

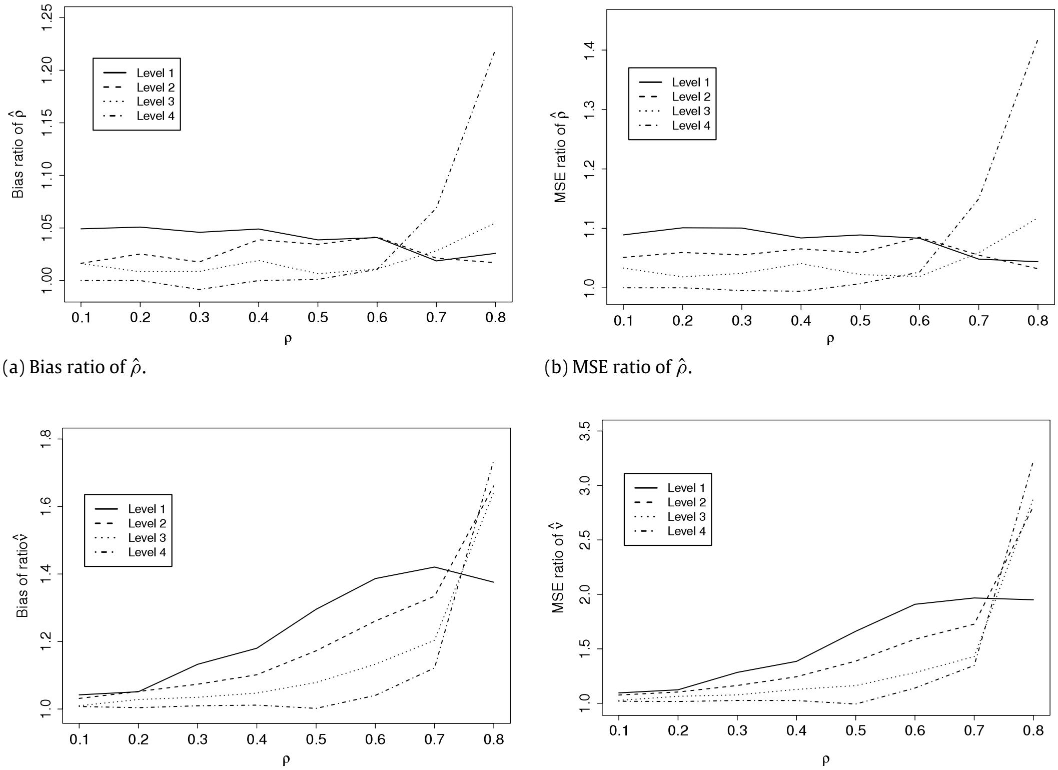
			SSP					
Degree of dependence			Low		Medium		High	
Model	Par.	Level	Bias	MSE	Bias	MSE	Bias	MSE
Gaussian	ρ	1	$1.09 \cdot 10^{-2}$	$1.84 \cdot 10^{-4}$	$8.47 \cdot 10^{-3}$	$1.12 \cdot 10^{-4}$	$3.98 \cdot 10^{-3}$	$2.45 \cdot 10^{-5}$
		2	$1.09 \cdot 10^{-2}$	$1.87 \cdot 10^{-4}$	$8.31 \cdot 10^{-3}$	$1.08 \cdot 10^{-4}$	$4.16 \cdot 10^{-3}$	$2.75 \cdot 10^{-5}$
		3	$1.09 \cdot 10^{-2}$	$1.86 \cdot 10^{-4}$	$8.82 \cdot 10^{-3}$	$1.20 \cdot 10^{-4}$	$4.27 \cdot 10^{-3}$	$2.90 \cdot 10^{-5}$
		4	$1.09 \cdot 10^{-2}$	$1.85 \cdot 10^{-4}$	$8.41 \cdot 10^{-3}$	$1.13 \cdot 10^{-4}$	$9.49 \cdot 10^{-3}$	$1.17 \cdot 10^{-4}$
		CPU	$1.63 \cdot 10^{-3}$		$2.98 \cdot 10^{-3}$		$1.67 \cdot 10^{-2}$	
Student's t	ρ	1	$1.24 \cdot 10^{-2}$	$2.40 \cdot 10^{-4}$	$9.67 \cdot 10^{-3}$	$1.47 \cdot 10^{-4}$	$4.76 \cdot 10^{-3}$	$3.57 \cdot 10^{-5}$
		2	$1.22 \cdot 10^{-2}$	$2.32 \cdot 10^{-4}$	$9.62 \cdot 10^{-3}$	$1.44 \cdot 10^{-4}$	$4.79 \cdot 10^{-3}$	$3.54 \cdot 10^{-5}$
		3	$1.20 \cdot 10^{-2}$	$2.22 \cdot 10^{-4}$	$9.32 \cdot 10^{-3}$	$1.38 \cdot 10^{-4}$	$5.00 \cdot 10^{-3}$	$3.99 \cdot 10^{-5}$
		4	$1.20 \cdot 10^{-2}$	$2.27 \cdot 10^{-4}$	$9.76 \cdot 10^{-3}$	$1.50 \cdot 10^{-4}$	$1.33 \cdot 10^{-2}$	$2.17 \cdot 10^{-4}$
	ν	1	$5.13 \cdot 10^{-1}$	$4.26 \cdot 10^{-1}$	$5.13 \cdot 10^{-1}$	$4.22 \cdot 10^{-1}$	$5.31 \cdot 10^{-1}$	$4.70 \cdot 10^{-1}$
		2	$5.06 \cdot 10^{-1}$	$4.15 \cdot 10^{-1}$	$5.03 \cdot 10^{-1}$	$4.26 \cdot 10^{-1}$	$5.15 \cdot 10^{-1}$	$4.38 \cdot 10^{-1}$
		3	$5.10 \cdot 10^{-1}$	$4.30 \cdot 10^{-1}$	$4.94 \cdot 10^{-1}$	$4.08 \cdot 10^{-1}$	$5.06 \cdot 10^{-1}$	$4.38 \cdot 10^{-1}$
		4	$5.11 \cdot 10^{-1}$	$4.35 \cdot 10^{-1}$	$5.18 \cdot 10^{-1}$	$4.47 \cdot 10^{-1}$	$5.23 \cdot 10^{-1}$	$4.68 \cdot 10^{-1}$
		CPU	$1.22 \cdot 10^{-2}$		$1.41 \cdot 10^{-2}$		$6.38 \cdot 10^{-3}$	
Clayton	α	1	$1.72 \cdot 10^{-2}$	$4.57 \cdot 10^{-4}$	$2.90 \cdot 10^{-2}$	$1.32 \cdot 10^{-3}$	$5.76 \cdot 10^{-2}$	$5.29 \cdot 10^{-3}$
		2	$1.66 \cdot 10^{-2}$	$4.40 \cdot 10^{-4}$	$2.77 \cdot 10^{-2}$	$1.21 \cdot 10^{-3}$	$5.72 \cdot 10^{-2}$	$5.06 \cdot 10^{-3}$
		3	$1.70 \cdot 10^{-2}$	$4.47 \cdot 10^{-4}$	$2.83 \cdot 10^{-2}$	$1.27 \cdot 10^{-3}$	$1.24 \cdot 10^{-1}$	$1.93 \cdot 10^{-2}$
		4	$1.67 \cdot 10^{-2}$	$4.48 \cdot 10^{-4}$	$2.81 \cdot 10^{-2}$	$1.29 \cdot 10^{-3}$	$6.96 \cdot 10^{-1}$	$4.93 \cdot 10^{-1}$
		CPU	$1.74 \cdot 10^{-3}$		$1.35 \cdot 10^{-3}$		$3.90 \cdot 10^{-4}$	

iterative transformations of the original pseudo observations. These transformations are functions of copulae at lower levels, and therefore depend on their parameters. The increasing estimator variance with the level number may indicate a stronger sensitivity to repeated transformations when the degree of dependence is high. Note that the SP and SSP estimates for the Gaussian PCC are virtually the same. This is as anticipated, since both estimators are semiparametrically efficient for that particular model [15]. Of course, the SSP estimator is asymptotically less efficient than the SP estimator for the Student's t and Clayton vines. This is reflected in higher bias and MSE. The difference is, however, mostly moderate to small, which indicates that the SSP estimator performs well relative to SP. Furthermore, the computing time of the former is much lower. The factor ranges from order 10^{-1} to 10^{-3} in favor of SSP. As one would expect, the most marked time gain is obtained for the Student's t vine, whose parameter vector is twice the size of the other two models'.

Figs. 3 and 4 show bias and MSE ratios of SSP versus SP for the parameter estimates of Model 2(b) and (c) as functions of the values of ρ and α, respectively, averaged over each of the four levels of the structure. These curves are based on additional simulations. More specifically, we repeated the experiments described above for Model 2(b) and (c) with the values $\{0.3, 0.4, 0.6, 0.7\}$ for ρ and $\{0.481, 0.71, 1.38, 1.95\}$ for α. The plots strengthen the impressions from Table 1. For

Table 2
Finite sample bias and MSE of $\boldsymbol{\theta}$ based on simulations of $n = 5000$ observations from Model 2(d) (see Fig. 2), i.e. consisting of different copula types, with medium ($\rho = 0.5$, $\alpha = 1$) dependence, as well as CPU time in seconds.

Copula	Par.	SP		SSP	
		Bias	MSE	Bias	MSE
12	ρ	$7.66 \cdot 10^{-3}$	$9.04 \cdot 10^{-5}$	$8.63 \cdot 10^{-3}$	$1.17 \cdot 10^{-4}$
23	ρ	$8.08 \cdot 10^{-3}$	$1.01 \cdot 10^{-4}$	$9.42 \cdot 10^{-3}$	$1.35 \cdot 10^{-4}$
	ν	$2.94 \cdot 10^{-1}$	$1.34 \cdot 10^{-1}$	$5.23 \cdot 10^{-1}$	$4.43 \cdot 10^{-1}$
34	α	$2.55 \cdot 10^{-2}$	$1.04 \cdot 10^{-3}$	$2.79 \cdot 10^{-2}$	$1.26 \cdot 10^{-3}$
45	ρ	$7.18 \cdot 10^{-3}$	$8.04 \cdot 10^{-5}$	$8.27 \cdot 10^{-3}$	$1.07 \cdot 10^{-4}$
13|2	ρ	$9.13 \cdot 10^{-3}$	$1.28 \cdot 10^{-4}$	$9.60 \cdot 10^{-3}$	$1.45 \cdot 10^{-4}$
	ν	$4.15 \cdot 10^{-1}$	$2.71 \cdot 10^{-1}$	$4.84 \cdot 10^{-1}$	$3.80 \cdot 10^{-1}$
24|3	α	$2.81 \cdot 10^{-2}$	$1.25 \cdot 10^{-3}$	$2.87 \cdot 10^{-2}$	$1.31 \cdot 10^{-3}$
35|4	ρ	$7.84 \cdot 10^{-3}$	$9.68 \cdot 10^{-5}$	$8.82 \cdot 10^{-3}$	$1.21 \cdot 10^{-4}$
14|23	ρ	$9.80 \cdot 10^{-3}$	$1.53 \cdot 10^{-4}$	$1.01 \cdot 10^{-2}$	$1.62 \cdot 10^{-4}$
	ν	$4.86 \cdot 10^{-1}$	$3.80 \cdot 10^{-1}$	$5.14 \cdot 10^{-1}$	$4.23 \cdot 10^{-1}$
25|34	α	$2.76 \cdot 10^{-2}$	$1.19 \cdot 10^{-3}$	$2.79 \cdot 10^{-2}$	$1.22 \cdot 10^{-3}$
15|234	ρ	$8.15 \cdot 10^{-3}$	$1.07 \cdot 10^{-4}$	$8.19 \cdot 10^{-3}$	$1.09 \cdot 10^{-4}$
CPU		1.06		$9.59 \cdot 10^{-3}$	



(a) Bias ratio of $\hat{\rho}$.

(b) MSE ratio of $\hat{\rho}$.

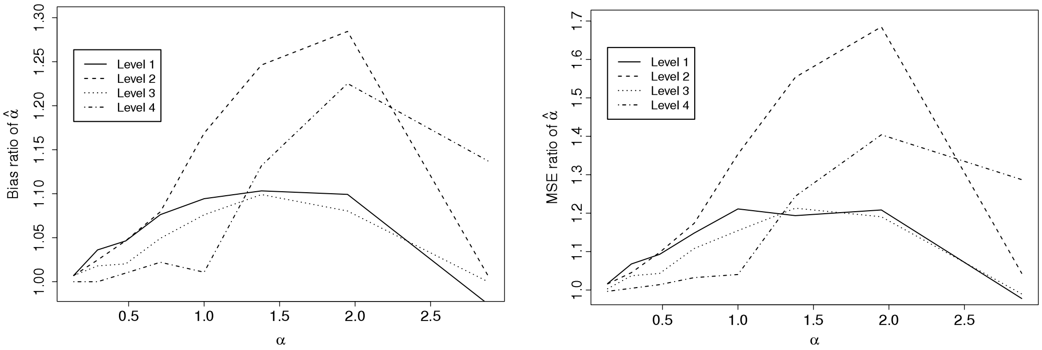
(c) Bias ratio of $\hat{\nu}$.

(d) MSE ratio of $\hat{\nu}$.

Fig. 3. Bias and MSE ratios, to the left and right, respectively, of the SSP (numerator) versus the SP (denominator) estimates $\hat{\rho}$ (top row) and $\hat{\nu}$ (bottom row) for Model 2(b) as functions of ρ, averaged over each of the four levels.

the Student's t vines (Model 2(b)), the bias and MSE ratios appear to increase with the degree of dependence at the top two (for ρ) or three (for ν) levels, whereas they decrease slightly at the ground level. This means that the SP estimator performs better relative to the SSP one with growing dependence, for the higher level parameters. The reason may be the earlier suggested stronger sensitivity to repeated transformations, which are likely to affect the SSP estimator more, due to its sequential nature. For the Clayton vines (Model 2(c)) on the other hand, the bias and MSE ratios seem to grow with the dependence up to a certain point, before they start decreasing again. There is no apparent reason for this different behavior. It could relate to characteristics of the Clayton copula, but may also be artifacts.

In order to study how the parameter estimation may affect dependence measures of interest, we also computed corresponding Kendall's τ and tail dependence coefficients. For the Student's t-copulae of Model 2(b), these two measures



(a) Bias ratio of $\hat{\alpha}$.

(b) MSE ratio of $\hat{\alpha}$.

Fig. 4. Bias and MSE ratios, to the left and right, respectively, of the SSP (numerator) versus the SP (denominator) estimates $\hat{\alpha}$ for Model 2(c) as functions of α, averaged over each of the four levels.

are given by $\tau = \frac{2}{\pi}\arcsin(\rho)$ and $\lambda = 2t_\nu\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right)$ (the upper and lower tail dependence coefficients are the same), respectively, where $t_\nu(\cdot)$ is the cdf of the Student's t-distribution with ν degrees of freedom. The lower tail dependence coefficients of the Clayton copulae of Model 2(c) are $\lambda_L = 2^{-\frac{1}{\alpha}}$, whereas $\tau = \frac{\alpha}{\alpha+2}$. The corresponding coefficient estimates were obtained by plugging the parameter estimates into these expressions. Table 3 shows the bias and MSE of these estimates. As one would expect, they follow the same pattern as for the original parameter estimates. The bias and MSE of τ decrease with the degree of dependence for both estimators, and so do those of λ_L for the Clayton copulae. The tail dependence coefficients of the Student's t-copulae appear to be mostly affected by the degrees of freedom estimates. Again, SSP performs worse than SP in terms of variance, and increasingly so with the degree of dependence. The difference between the two estimators is, however, moderate.

For the mixed vine of Model 2(d) (Table 2), the results are rather similar. The overall degree of dependence is medium. As expected, the bias and MSE are rather constant over the different levels for copulae of the same type. Moreover, the performance of the two estimators for the Clayton and Student's t-copulae of Model 2(d) are virtually the same as for the corresponding levels in Model 2(b) and (c), respectively. Since the full model no longer is a five-dimensional Gaussian copula, neither estimator is semiparametrically efficient, so the bias and MSE for the Gaussian copulae are higher than the ones reported in Table 1. Again the SSP estimator performs quite well compared to the SP estimator in terms of efficiency, while its CPU time is drastically lower.

4.2. Finite and small sample size

The experiments in Section 4.1 were based on a large number of observations. To see how the two estimators perform on samples of smaller size, we repeated the simulations from Model 2(a)–(c), reducing n, first to 500, and then to 50. The results from the simulations with medium dependence are summarized in Table 4. As the sample size decreases, we expect the estimators' variance to increase and finally explode. For $n = 500$, the finite sample bias and MSE of the correlations ρ and the parameters α are actually not discouragingly high. Estimation of the degrees of freedom parameters ν apparently requires a larger sample size. As n decreases to 50, none of the parameters are well estimated. In particular, the two estimators struggle with ν, especially SSP. Once more, the efficiency of the SSP estimator for ρ and α is rather good, compared to the SP estimator. Actually, for these parameters, the SSP estimator appears to suffer slightly less under reduced sample size than its competitor. That could relate to the fact that the numerical optimizations are performed for each copula individually, as opposed to all at once. Also, the computing time of the former is lower, especially for $n = 50$.

4.3. Robustness

Since they are based on the empirical margins, both the SP and the SSP estimators are robust to deviations from the chosen marginal distributions. However, their properties do rely on the correctness of the specified dependence structure.

As described in [1], selecting the pair-copulae constituting the PCC for a given dataset is a levelwise procedure. At the ground level, it consists in a set of bivariate copula selection problems, that may be handled, for instance, with copula goodness-of-fit tests [13,4]. However, as one proceeds into the structure, one must condition on the pair-copulae chosen at preceding levels to be able to compute the necessary conditional distributions, that are pair-copula arguments (as explained in Section 3). One would therefore expect the difficulty of finding adequate copulae to increase with the level number.

We wish to explore how the two estimators perform when the specified model deviates from the true model. In order to do that, we have perturbed the original models (from Section 4.1). More specifically, we have mixed the copulae of each of the two models 2(b) and (c) with the copulae of Model 2(a), using the same degree of dependence and a fixed mixing probability p. For instance, the pair (u_1, u_2) of observations is drawn from a Student's t-copula (Clayton copula) with probability $1 - p$

Table 3
Finite sample bias and MSE of Kendall's τ and (lower) tail dependence coefficients λ (λ_L) based on simulations of $n = 5000$ observations from Model 2(b) and (c) (see Fig. 2), i.e. consisting of only one copula type, with low ($\rho = 0.2, \alpha = 0.294$), medium ($\rho = 0.5, \alpha = 1$) and high ($\rho = 0.8, \alpha = 2.88$) dependence.

			SP					
Degree of dependence			Low		Medium		High	
Model	Par.	Level	Bias	MSE	Bias	MSE	Bias	MSE
Student's t	τ	1	$7.69 \cdot 10^{-3}$	$9.21 \cdot 10^{-5}$	$6.85 \cdot 10^{-3}$	$7.32 \cdot 10^{-5}$	$4.93 \cdot 10^{-3}$	$3.86 \cdot 10^{-5}$
		2	$7.71 \cdot 10^{-3}$	$9.26 \cdot 10^{-5}$	$6.84 \cdot 10^{-3}$	$7.33 \cdot 10^{-5}$	$5.00 \cdot 10^{-3}$	$3.87 \cdot 10^{-5}$
		3	$7.72 \cdot 10^{-3}$	$9.21 \cdot 10^{-5}$	$6.81 \cdot 10^{-3}$	$7.31 \cdot 10^{-5}$	$5.01 \cdot 10^{-3}$	$3.97 \cdot 10^{-5}$
		4	$7.80 \cdot 10^{-3}$	$9.61 \cdot 10^{-5}$	$7.16 \cdot 10^{-3}$	$8.07 \cdot 10^{-5}$	$1.14 \cdot 10^{-2}$	$1.66 \cdot 10^{-4}$
	λ	1	$1.18 \cdot 10^{-2}$	$2.19 \cdot 10^{-4}$	$1.40 \cdot 10^{-2}$	$3.05 \cdot 10^{-4}$	$1.43 \cdot 10^{-2}$	$3.26 \cdot 10^{-4}$
		2	$1.15 \cdot 10^{-2}$	$2.09 \cdot 10^{-4}$	$1.49 \cdot 10^{-2}$	$3.55 \cdot 10^{-4}$	$1.22 \cdot 10^{-2}$	$2.33 \cdot 10^{-4}$
		3	$1.16 \cdot 10^{-2}$	$2.10 \cdot 10^{-4}$	$1.56 \cdot 10^{-2}$	$3.86 \cdot 10^{-4}$	$1.30 \cdot 10^{-2}$	$2.61 \cdot 10^{-4}$
		4	$1.19 \cdot 10^{-2}$	$2.26 \cdot 10^{-4}$	$1.74 \cdot 10^{-2}$	$4.80 \cdot 10^{-4}$	$1.83 \cdot 10^{-2}$	$4.90 \cdot 10^{-4}$
Clayton	τ	1	$6.29 \cdot 10^{-3}$	$6.14 \cdot 10^{-5}$	$5.89 \cdot 10^{-3}$	$5.38 \cdot 10^{-5}$	$5.04 \cdot 10^{-3}$	$3.97 \cdot 10^{-5}$
		2	$6.15 \cdot 10^{-3}$	$6.04 \cdot 10^{-5}$	$5.28 \cdot 10^{-3}$	$4.42 \cdot 10^{-5}$	$4.84 \cdot 10^{-3}$	$3.56 \cdot 10^{-5}$
		3	$6.34 \cdot 10^{-3}$	$6.22 \cdot 10^{-5}$	$5.85 \cdot 10^{-3}$	$5.45 \cdot 10^{-5}$	$1.08 \cdot 10^{-2}$	$1.49 \cdot 10^{-4}$
		4	$6.33 \cdot 10^{-3}$	$6.43 \cdot 10^{-5}$	$6.21 \cdot 10^{-3}$	$6.19 \cdot 10^{-5}$	$5.90 \cdot 10^{-2}$	$3.58 \cdot 10^{-3}$
	λ_L	1	$1.26 \cdot 10^{-2}$	$2.46 \cdot 10^{-4}$	$9.18 \cdot 10^{-3}$	$1.31 \cdot 10^{-4}$	$3.97 \cdot 10^{-3}$	$2.48 \cdot 10^{-5}$
		2	$1.23 \cdot 10^{-2}$	$2.42 \cdot 10^{-4}$	$8.24 \cdot 10^{-3}$	$1.08 \cdot 10^{-4}$	$3.82 \cdot 10^{-3}$	$2.22 \cdot 10^{-5}$
		3	$1.26 \cdot 10^{-2}$	$2.46 \cdot 10^{-4}$	$9.14 \cdot 10^{-3}$	$1.33 \cdot 10^{-4}$	$8.59 \cdot 10^{-3}$	$9.44 \cdot 10^{-5}$
		4	$1.26 \cdot 10^{-2}$	$2.55 \cdot 10^{-4}$	$9.72 \cdot 10^{-3}$	$1.52 \cdot 10^{-4}$	$4.98 \cdot 10^{-2}$	$2.56 \cdot 10^{-3}$

			SSP					
Degree of dependence			Low		Medium		High	
Model	Par.	Level	Bias	MSE	Bias	MSE	Bias	MSE
Student's t	τ	1	$8.08 \cdot 10^{-3}$	$1.01 \cdot 10^{-4}$	$7.11 \cdot 10^{-3}$	$7.94 \cdot 10^{-5}$	$5.06 \cdot 10^{-3}$	$4.03 \cdot 10^{-5}$
		2	$7.94 \cdot 10^{-3}$	$9.80 \cdot 10^{-5}$	$7.07 \cdot 10^{-3}$	$7.80 \cdot 10^{-5}$	$5.09 \cdot 10^{-3}$	$3.98 \cdot 10^{-5}$
		3	$7.77 \cdot 10^{-3}$	$9.39 \cdot 10^{-5}$	$6.85 \cdot 10^{-3}$	$7.44 \cdot 10^{-5}$	$5.28 \cdot 10^{-3}$	$4.43 \cdot 10^{-5}$
		4	$7.80 \cdot 10^{-3}$	$9.61 \cdot 10^{-5}$	$7.17 \cdot 10^{-3}$	$8.08 \cdot 10^{-5}$	$1.39 \cdot 10^{-2}$	$2.34 \cdot 10^{-4}$
	λ	1	$1.23 \cdot 10^{-2}$	$2.41 \cdot 10^{-4}$	$1.76 \cdot 10^{-2}$	$4.82 \cdot 10^{-4}$	$1.92 \cdot 10^{-2}$	$5.86 \cdot 10^{-4}$
		2	$1.20 \cdot 10^{-2}$	$2.26 \cdot 10^{-4}$	$1.70 \cdot 10^{-2}$	$4.61 \cdot 10^{-4}$	$1.84 \cdot 10^{-2}$	$5.34 \cdot 10^{-4}$
		3	$1.18 \cdot 10^{-2}$	$2.19 \cdot 10^{-4}$	$1.67 \cdot 10^{-2}$	$4.37 \cdot 10^{-4}$	$1.84 \cdot 10^{-2}$	$5.52 \cdot 10^{-4}$
		4	$1.19 \cdot 10^{-2}$	$2.24 \cdot 10^{-4}$	$1.74 \cdot 10^{-2}$	$4.73 \cdot 10^{-4}$	$2.55 \cdot 10^{-2}$	$1.02 \cdot 10^{-3}$
Clayton	τ	1	$6.51 \cdot 10^{-3}$	$6.56 \cdot 10^{-5}$	$6.43 \cdot 10^{-3}$	$6.49 \cdot 10^{-5}$	$4.83 \cdot 10^{-3}$	$3.73 \cdot 10^{-5}$
		2	$6.28 \cdot 10^{-3}$	$6.31 \cdot 10^{-5}$	$6.17 \cdot 10^{-3}$	$6.01 \cdot 10^{-5}$	$4.84 \cdot 10^{-3}$	$3.64 \cdot 10^{-5}$
		3	$6.45 \cdot 10^{-3}$	$6.45 \cdot 10^{-5}$	$6.30 \cdot 10^{-3}$	$6.32 \cdot 10^{-5}$	$1.07 \cdot 10^{-2}$	$1.46 \cdot 10^{-4}$
		4	$6.34 \cdot 10^{-3}$	$6.45 \cdot 10^{-5}$	$6.28 \cdot 10^{-3}$	$6.48 \cdot 10^{-5}$	$6.84 \cdot 10^{-2}$	$4.79 \cdot 10^{-3}$
	λ_L	1	$1.30 \cdot 10^{-2}$	$2.63 \cdot 10^{-4}$	$1.00 \cdot 10^{-2}$	$1.57 \cdot 10^{-4}$	$3.78 \cdot 10^{-3}$	$2.28 \cdot 10^{-5}$
		2	$1.25 \cdot 10^{-2}$	$2.52 \cdot 10^{-4}$	$9.63 \cdot 10^{-3}$	$1.47 \cdot 10^{-4}$	$3.80 \cdot 10^{-3}$	$2.25 \cdot 10^{-5}$
		3	$1.28 \cdot 10^{-2}$	$2.55 \cdot 10^{-4}$	$9.85 \cdot 10^{-3}$	$1.55 \cdot 10^{-4}$	$8.52 \cdot 10^{-3}$	$9.31 \cdot 10^{-5}$
		4	$1.26 \cdot 10^{-2}$	$2.55 \cdot 10^{-4}$	$9.86 \cdot 10^{-3}$	$1.61 \cdot 10^{-4}$	$5.84 \cdot 10^{-2}$	$3.50 \cdot 10^{-3}$

and from a Gaussian copula with probability p. In all experiments, we used $n = 5000$ and medium dependence, letting the mixing probability take each of the values $\{0.05, 0.1, 0.2, 1\}$. The first three values of p correspond to rather small or moderate misspecifications of the model, that one may expect if the amount of data is acceptable and the copulae are soundly chosen. When $p = 1$, the true model is a Gaussian D-vine (Model 2(a)), but we assume and estimate a Student's t (Model 2(b)) or a Clayton (Model 2(c)). The latter case, i.e. fitting Clayton copulae to Gaussian data, may be somewhat far-fetched, especially in the lower levels, but we include it for comparison.

The results are summarized by levels in Tables 5 and 6. Since the estimated models have different parameters from the true model, we cannot compare the estimates and true parameters directly. Instead, we have computed the finite sample bias and MSE of the corresponding Kendall's τ and (lower) tail dependence coefficients. Further, these experiments were run on a different computer than the others. More specifically, this computer has 64 GB RAM and an X7542 @ 2.67 GHz processor with 6 CPU kernels, and is therefore faster. Hence, the reported CPU times in Tables 5 and 6 cannot be directly compared to the ones in Sections 4.1 and 4.2.

Both the SP and the SSP estimators of the Kendall's τ coefficients of Model 2(b) perform almost as well as in the non-perturbed case (Table 3), even for the fully misspecified model ($p = 1$). The reason for this is probably that the Student's t-copulae and the Gaussian ones, that we have mixed with, have the same correlations ($\rho = 0.5$). The corresponding tail dependence estimates are, however, decreasingly accurate as the mixing probability grows, as one would expect. That is also the case for the Kendall's τ and lower tail dependence coefficients of the Clayton copulae of Model 2(c). Both estimators perform worse for this model, probably because the contrast between the asymmetric Clayton copulae and the symmetric Gaussian ones is larger than between Gaussian and Student's t-copulae. Moreover, the computing times of both estimators grow with the mixing probability, due to the need for extra iterations before convergence.

Table 4
Finite sample bias and MSE of $\boldsymbol{\theta}$ based on simulations of $n = 500$ and $n = 50$ observations from Model 2(a)–(c) (see Fig. 2), i.e. consisting of only one copula type, with medium ($\rho = 0.5$, $\alpha = 1$) dependence, as well as CPU time in seconds.

			SP			
Sample size			$n = 500$		$n = 50$	
Model	Par.	Level	Bias	MSE	Bias	MSE
Gaussian	ρ	1	$2.73 \cdot 10^{-2}$	$1.15 \cdot 10^{-3}$	$9.09 \cdot 10^{-2}$	$1.25 \cdot 10^{-2}$
		2	$2.71 \cdot 10^{-2}$	$1.16 \cdot 10^{-3}$	$9.10 \cdot 10^{-2}$	$1.30 \cdot 10^{-2}$
		3	$2.69 \cdot 10^{-2}$	$1.15 \cdot 10^{-3}$	$9.24 \cdot 10^{-2}$	$1.35 \cdot 10^{-2}$
		4	$2.68 \cdot 10^{-2}$	$1.16 \cdot 10^{-3}$	$9.75 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$
		CPU	$2.10 \cdot 10^{-4}$		$1.70 \cdot 10^{-4}$	
Student's t	ρ	1	$3.00 \cdot 10^{-2}$	$1.43 \cdot 10^{-3}$	$9.90 \cdot 10^{-2}$	$1.50 \cdot 10^{-2}$
		2	$3.04 \cdot 10^{-2}$	$1.44 \cdot 10^{-3}$	$9.82 \cdot 10^{-2}$	$1.52 \cdot 10^{-2}$
		3	$2.99 \cdot 10^{-2}$	$1.42 \cdot 10^{-3}$	$1.04 \cdot 10^{-1}$	$1.69 \cdot 10^{-2}$
		4	$3.13 \cdot 10^{-2}$	$1.53 \cdot 10^{-3}$	$1.05 \cdot 10^{-1}$	$1.78 \cdot 10^{-2}$
	ν	1	1.48	4.91	20.3	2030
		2	1.56	6.03	24.6	2430
		3	1.79	8.57	27.9	2930
		4	2.10	12.0	30.0	3020
		CPU	$9.44 \cdot 10^{-3}$		$1.35 \cdot 10^{-3}$	
Clayton	α	1	$8.21 \cdot 10^{-2}$	$1.06 \cdot 10^{-2}$	$2.93 \cdot 10^{-1}$	$1.42 \cdot 10^{-1}$
		2	$7.97 \cdot 10^{-2}$	$1.00 \cdot 10^{-2}$	$2.75 \cdot 10^{-1}$	$1.25 \cdot 10^{-1}$
		3	$8.46 \cdot 10^{-2}$	$1.09 \cdot 10^{-2}$	$2.88 \cdot 10^{-1}$	$1.33 \cdot 10^{-1}$
		4	$9.41 \cdot 10^{-2}$	$1.31 \cdot 10^{-2}$	$3.41 \cdot 10^{-1}$	$1.81 \cdot 10^{-1}$
		CPU	$4.50 \cdot 10^{-4}$		$1.00 \cdot 10^{-5}$	
			SSP			
Sample size			$n = 500$		$n = 50$	
Model	Par.	Level	Bias	MSE	Bias	MSE
Gaussian	ρ	1	$2.72 \cdot 10^{-2}$	$1.16 \cdot 10^{-3}$	$9.18 \cdot 10^{-2}$	$1.30 \cdot 10^{-2}$
		2	$2.72 \cdot 10^{-2}$	$1.16 \cdot 10^{-3}$	$9.16 \cdot 10^{-2}$	$1.31 \cdot 10^{-2}$
		3	$2.69 \cdot 10^{-2}$	$1.15 \cdot 10^{-3}$	$9.22 \cdot 10^{-2}$	$1.34 \cdot 10^{-2}$
		4	$2.68 \cdot 10^{-2}$	$1.16 \cdot 10^{-3}$	$9.74 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$
		CPU	$2.00 \cdot 10^{-5}$		$1.00 \cdot 10^{-5}$	
Student's t	ρ	1	$3.13 \cdot 10^{-2}$	$1.54 \cdot 10^{-3}$	$1.01 \cdot 10^{-1}$	$1.56 \cdot 10^{-2}$
		2	$3.11 \cdot 10^{-2}$	$1.51 \cdot 10^{-3}$	$9.88 \cdot 10^{-2}$	$1.52 \cdot 10^{-2}$
		3	$3.04 \cdot 10^{-2}$	$1.46 \cdot 10^{-3}$	$1.03 \cdot 10^{-1}$	$1.67 \cdot 10^{-2}$
		4	$3.16 \cdot 10^{-2}$	$1.54 \cdot 10^{-3}$	$1.04 \cdot 10^{-1}$	$1.86 \cdot 10^{-2}$
	ν	1	2.60	142.00	89.0	24900
		2	2.35	82.2	97.0	27100
		3	2.98	203.00	111.00	31300
		4	2.62	52.7	127.00	35900
		CPU	$3.00 \cdot 10^{-5}$		$<1.00 \cdot 10^{-18}$	
Clayton	α	1	$9.01 \cdot 10^{-2}$	$1.28 \cdot 10^{-2}$	$3.23 \cdot 10^{-1}$	$1.74 \cdot 10^{-1}$
		2	$9.03 \cdot 10^{-2}$	$1.27 \cdot 10^{-2}$	$2.83 \cdot 10^{-1}$	$1.30 \cdot 10^{-1}$
		3	$8.97 \cdot 10^{-2}$	$1.24 \cdot 10^{-2}$	$2.88 \cdot 10^{-1}$	$1.27 \cdot 10^{-1}$
		4	$1.01 \cdot 10^{-1}$	$1.52 \cdot 10^{-2}$	$3.55 \cdot 10^{-1}$	$1.79 \cdot 10^{-1}$
		CPU	$3.00 \cdot 10^{-5}$		$1.00 \cdot 10^{-5}$	

Fig. 5 shows bias ratios of SSP versus SP for the Kendall's τ and tail dependence coefficients as functions of the mixing probability p. For the perturbed Student's t-vine, these ratios tend to decrease with the mixing probability, which means that the SSP estimator performs relatively better compared to SP. The ratios for the perturbed Clayton model are less conclusive. The SP estimator is superior for λ_L at the ground level and τ at the upper two levels. Otherwise, the SSP estimator performs just as well as, or even better than SP. All in all, the difference between the performance of the SP and the SSP estimators is reduced with an increasing degree of perturbation, maybe because the latter uses information only from preceding levels, and not from the following. Hence, when the model assumptions are not completely accurate, the gain from using the SP estimator seems to be smaller.

4.4. Large dimension d

As mentioned earlier, the SP estimator is computationally too demanding and time consuming for high-dimensional problems. Our belief is that the SSP might be a good alternative in such cases. We have therefore tried it on a 50-dimensional

Table 5
Finite sample bias and MSE of Kendall's τ and (lower) tail dependence coefficients λ (λ_L) based on simulations of $n = 5000$ observations from Model 2(b) and (c) mixed with Model 2(a) (see Fig. 2), with medium dependence, for different mixing probabilities p, as well as CPU time in seconds.

			SP					
Mixing probability			$p = 0.05$		$p = 0.1$		$p = 0.2$	
Model	Par.	Level	Bias	MSE	Bias	MSE	Bias	MSE
Student's t	τ	1	$6.75 \cdot 10^{-3}$	$7.12 \cdot 10^{-5}$	$6.90 \cdot 10^{-3}$	$7.39 \cdot 10^{-5}$	$6.80 \cdot 10^{-3}$	$7.27 \cdot 10^{-5}$
		2	$6.95 \cdot 10^{-3}$	$7.55 \cdot 10^{-5}$	$6.93 \cdot 10^{-3}$	$7.60 \cdot 10^{-5}$	$6.90 \cdot 10^{-3}$	$7.37 \cdot 10^{-5}$
		3	$6.82 \cdot 10^{-3}$	$7.28 \cdot 10^{-5}$	$6.98 \cdot 10^{-3}$	$7.71 \cdot 10^{-5}$	$7.03 \cdot 10^{-3}$	$7.70 \cdot 10^{-5}$
		4	$7.54 \cdot 10^{-3}$	$8.60 \cdot 10^{-5}$	$7.27 \cdot 10^{-3}$	$8.19 \cdot 10^{-5}$	$7.25 \cdot 10^{-3}$	$8.36 \cdot 10^{-5}$
	λ	1	$1.45 \cdot 10^{-2}$	$3.29 \cdot 10^{-4}$	$1.50 \cdot 10^{-2}$	$3.57 \cdot 10^{-4}$	$1.81 \cdot 10^{-2}$	$5.14 \cdot 10^{-4}$
		2	$1.55 \cdot 10^{-2}$	$3.82 \cdot 10^{-4}$	$1.59 \cdot 10^{-2}$	$4.00 \cdot 10^{-4}$	$1.73 \cdot 10^{-2}$	$4.70 \cdot 10^{-4}$
		3	$1.69 \cdot 10^{-2}$	$4.60 \cdot 10^{-4}$	$1.72 \cdot 10^{-2}$	$4.67 \cdot 10^{-4}$	$1.76 \cdot 10^{-2}$	$4.88 \cdot 10^{-4}$
		4	$1.73 \cdot 10^{-2}$	$4.77 \cdot 10^{-4}$	$1.78 \cdot 10^{-2}$	$5.13 \cdot 10^{-4}$	$1.86 \cdot 10^{-2}$	$5.62 \cdot 10^{-4}$
		CPU	$1.84 \cdot 10^{-1}$		$2.22 \cdot 10^{-1}$		$8.03 \cdot 10^{-1}$	
Clayton	τ	1	$1.11 \cdot 10^{-2}$	$1.71 \cdot 10^{-4}$	$2.03 \cdot 10^{-2}$	$4.76 \cdot 10^{-4}$	$3.89 \cdot 10^{-2}$	$1.59 \cdot 10^{-3}$
		2	$8.62 \cdot 10^{-3}$	$1.09 \cdot 10^{-4}$	$1.48 \cdot 10^{-2}$	$2.72 \cdot 10^{-4}$	$2.80 \cdot 10^{-2}$	$8.48 \cdot 10^{-4}$
		3	$9.83 \cdot 10^{-3}$	$1.38 \cdot 10^{-4}$	$1.70 \cdot 10^{-2}$	$3.52 \cdot 10^{-4}$	$3.13 \cdot 10^{-2}$	$1.05 \cdot 10^{-3}$
		4	$1.52 \cdot 10^{-2}$	$2.98 \cdot 10^{-4}$	$2.68 \cdot 10^{-2}$	$7.97 \cdot 10^{-4}$	$4.65 \cdot 10^{-2}$	$2.26 \cdot 10^{-3}$
	λ_L	1	$1.23 \cdot 10^{-2}$	$2.28 \cdot 10^{-4}$	$1.86 \cdot 10^{-2}$	$4.77 \cdot 10^{-4}$	$3.57 \cdot 10^{-2}$	$1.50 \cdot 10^{-3}$
		2	$1.44 \cdot 10^{-2}$	$2.94 \cdot 10^{-4}$	$2.66 \cdot 10^{-2}$	$8.51 \cdot 10^{-4}$	$5.43 \cdot 10^{-2}$	$3.14 \cdot 10^{-3}$
		3	$1.32 \cdot 10^{-2}$	$2.59 \cdot 10^{-4}$	$2.33 \cdot 10^{-2}$	$6.97 \cdot 10^{-4}$	$4.87 \cdot 10^{-2}$	$2.58 \cdot 10^{-3}$
		4	$1.12 \cdot 10^{-2}$	$1.99 \cdot 10^{-4}$	$1.34 \cdot 10^{-2}$	$2.71 \cdot 10^{-4}$	$2.40 \cdot 10^{-2}$	$8.12 \cdot 10^{-4}$
		CPU	$3.30 \cdot 10^{-2}$		$1.13 \cdot 10^{-3}$		$3.30 \cdot 10^{-2}$	

			SSP					
Mixing probability			$p = 0.05$		$p = 0.1$		$p = 0.2$	
Model	Par.	Level	Bias	MSE	Bias	MSE	Bias	MSE
Student's t	τ	1	$7.02 \cdot 10^{-3}$	$7.70 \cdot 10^{-5}$	$7.09 \cdot 10^{-3}$	$7.81 \cdot 10^{-5}$	$6.96 \cdot 10^{-3}$	$7.63 \cdot 10^{-5}$
		2	$7.15 \cdot 10^{-3}$	$8.00 \cdot 10^{-5}$	$7.16 \cdot 10^{-3}$	$8.03 \cdot 10^{-5}$	$7.08 \cdot 10^{-3}$	$7.76 \cdot 10^{-5}$
		3	$6.90 \cdot 10^{-3}$	$7.46 \cdot 10^{-5}$	$7.06 \cdot 10^{-3}$	$7.87 \cdot 10^{-5}$	$7.07 \cdot 10^{-3}$	$7.79 \cdot 10^{-5}$
		4	$7.59 \cdot 10^{-3}$	$8.68 \cdot 10^{-5}$	$7.29 \cdot 10^{-3}$	$8.24 \cdot 10^{-5}$	$7.26 \cdot 10^{-3}$	$8.42 \cdot 10^{-5}$
	λ	1	$1.84 \cdot 10^{-2}$	$5.20 \cdot 10^{-4}$	$1.82 \cdot 10^{-2}$	$5.22 \cdot 10^{-4}$	$1.98 \cdot 10^{-2}$	$6.23 \cdot 10^{-4}$
		2	$1.77 \cdot 10^{-2}$	$4.94 \cdot 10^{-4}$	$1.81 \cdot 10^{-2}$	$5.15 \cdot 10^{-4}$	$1.90 \cdot 10^{-2}$	$5.66 \cdot 10^{-4}$
		3	$1.79 \cdot 10^{-2}$	$5.12 \cdot 10^{-4}$	$1.82 \cdot 10^{-2}$	$5.18 \cdot 10^{-4}$	$1.89 \cdot 10^{-2}$	$5.54 \cdot 10^{-4}$
		4	$1.71 \cdot 10^{-2}$	$4.64 \cdot 10^{-4}$	$1.77 \cdot 10^{-2}$	$5.04 \cdot 10^{-4}$	$1.87 \cdot 10^{-2}$	$5.63 \cdot 10^{-4}$
		CPU	$1.22 \cdot 10^{-3}$		$1.60 \cdot 10^{-3}$		$5.71 \cdot 10^{-3}$	
Clayton	τ	1	$7.10 \cdot 10^{-3}$	$8.01 \cdot 10^{-5}$	$9.49 \cdot 10^{-3}$	$1.33 \cdot 10^{-4}$	$1.73 \cdot 10^{-2}$	$3.58 \cdot 10^{-4}$
		2	$8.65 \cdot 10^{-3}$	$1.12 \cdot 10^{-4}$	$1.41 \cdot 10^{-2}$	$2.56 \cdot 10^{-4}$	$2.68 \cdot 10^{-2}$	$7.82 \cdot 10^{-4}$
		3	$1.19 \cdot 10^{-2}$	$1.94 \cdot 10^{-4}$	$2.13 \cdot 10^{-2}$	$5.23 \cdot 10^{-4}$	$3.94 \cdot 10^{-2}$	$1.62 \cdot 10^{-3}$
		4	$1.88 \cdot 10^{-2}$	$4.28 \cdot 10^{-4}$	$3.40 \cdot 10^{-2}$	$1.24 \cdot 10^{-3}$	$6.13 \cdot 10^{-2}$	$3.86 \cdot 10^{-3}$
	λ_L	1	$1.98 \cdot 10^{-2}$	$5.23 \cdot 10^{-4}$	$3.69 \cdot 10^{-2}$	$1.53 \cdot 10^{-3}$	$7.23 \cdot 10^{-2}$	$5.39 \cdot 10^{-3}$
		2	$1.53 \cdot 10^{-2}$	$3.34 \cdot 10^{-4}$	$2.80 \cdot 10^{-2}$	$9.46 \cdot 10^{-4}$	$5.64 \cdot 10^{-2}$	$3.36 \cdot 10^{-3}$
		3	$1.17 \cdot 10^{-2}$	$2.11 \cdot 10^{-4}$	$1.75 \cdot 10^{-2}$	$4.37 \cdot 10^{-4}$	$3.48 \cdot 10^{-2}$	$1.43 \cdot 10^{-3}$
		4	$1.22 \cdot 10^{-2}$	$2.36 \cdot 10^{-4}$	$1.35 \cdot 10^{-2}$	$2.84 \cdot 10^{-4}$	$1.55 \cdot 10^{-2}$	$3.76 \cdot 10^{-4}$
		CPU	$1.13 \cdot 10^{-3}$		$3.30 \cdot 10^{-2}$		$1.13 \cdot 10^{-3}$	

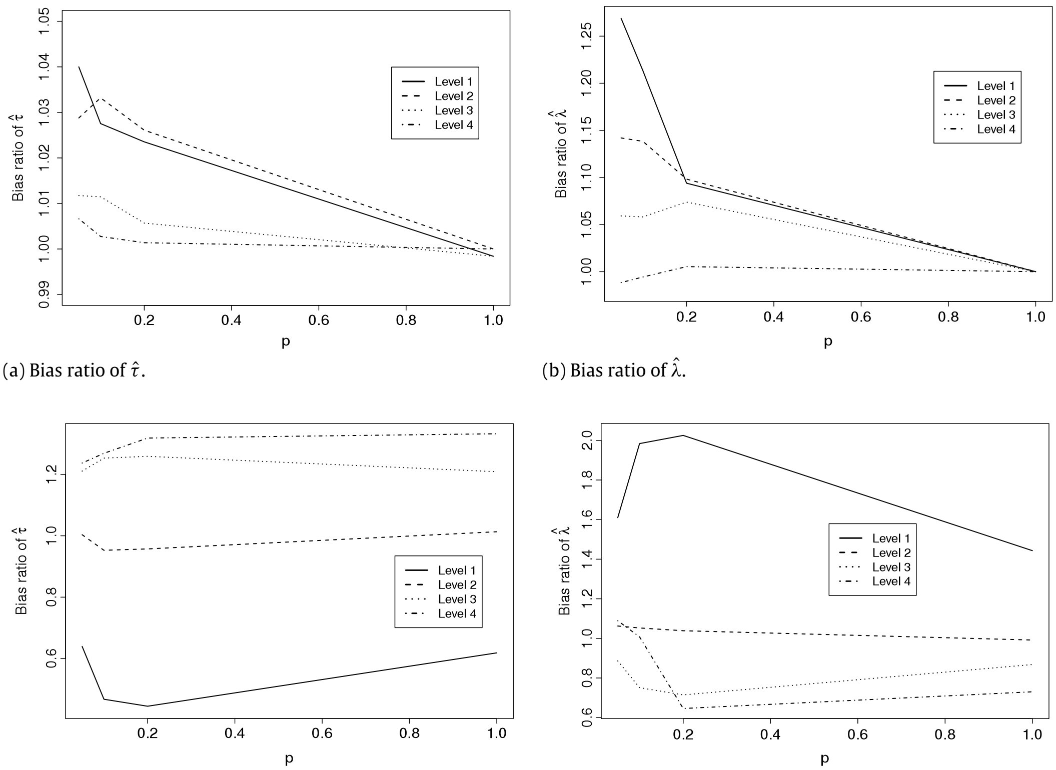
D-vine of Student's copulae with $\rho = 0.2$, corresponding to low dependence. This model has as many as 2450 parameters. Optimizing over so many parameters simultaneously would not only be highly time consuming, but also numerically dubious. We have therefore only considered the SSP estimator for this model. Moreover, we let $n = 5000$ and $N = 1000$ as in the experiments of Section 4.1.

Fig. 6 displays the bias and MSE of the parameter estimates $\hat{\rho}$ and $\hat{\nu}$, averaged over each level. These are rather low up to level 20 for $\hat{\rho}$ and up to level 30 for $\hat{\nu}$, after which they explode. This is due to numerical problems with the repeated transformations of the original data. After a certain level, the computed estimates in practice tend toward independence, i.e. partial correlations of approximately 0 and a high number of degrees of freedom. The subsequent decrease (and increase for $\hat{\rho}$) of the bias and MSE seems unintuitive. The correlation estimates $\hat{\rho}$ appear to fluctuate around 0, whereas the estimates $\hat{\nu}$ are almost halved on average. Actually, there is not that much difference between a bivariate t-copula with 100 and 200 degrees of freedom. Hence, this may just be an artifact. Note that what we are trying to estimate at these levels, is conditional dependence with a very high number of conditioning variables. It is not that surprising that it is difficult to estimate such higher order dependences. The question is of course, how different is the estimated distribution from the true distribution?

In a pair-copula construction of dimension d, only $d - 1$ of the pairwise dependences are modeled unconditionally. For instance, in this 50-dimensional D-vine, the dependence between U_1 and U_{40} is modeled through $C_{1,40|2...39}$, i.e. conditionally on 38 variables, as well as conditionals in lower levels, involving a subset of the 40 variables in question. In most applications, it is the unconditional dependence, or at least lower order dependences, one is interested in. As long as the bottom levels

Table 6
Finite sample bias and MSE of Kendall's τ and (lower) tail dependence coefficients λ (λ_L) based on simulations of $n = 5000$ observations from Model 2(a), assuming Model 2(b) and (c), respectively, (see Fig. 2), as well as CPU time in seconds.

Model	Par.	Level	SP		SSP	
			Bias	MSE	Bias	MSE
Assuming Student's t	ρ	1	$6.25 \cdot 10^{-3}$	$6.13 \cdot 10^{-5}$	$6.24 \cdot 10^{-3}$	$6.12 \cdot 10^{-5}$
		2	$6.16 \cdot 10^{-3}$	$6.01 \cdot 10^{-5}$	$6.16 \cdot 10^{-3}$	$6.01 \cdot 10^{-5}$
		3	$6.26 \cdot 10^{-3}$	$6.09 \cdot 10^{-5}$	$6.25 \cdot 10^{-3}$	$6.09 \cdot 10^{-5}$
		4	$6.05 \cdot 10^{-3}$	$5.87 \cdot 10^{-5}$	$6.05 \cdot 10^{-3}$	$5.87 \cdot 10^{-5}$
	ν	1	$1.42 \cdot 10^{-1}$	$2.00 \cdot 10^{-2}$	$1.41 \cdot 10^{-1}$	$2.00 \cdot 10^{-2}$
		2	$1.37 \cdot 10^{-1}$	$1.83 \cdot 10^{-2}$	$1.37 \cdot 10^{-1}$	$1.83 \cdot 10^{-2}$
		3	$1.39 \cdot 10^{-1}$	$1.91 \cdot 10^{-2}$	$1.38 \cdot 10^{-1}$	$1.91 \cdot 10^{-2}$
		4	$1.43 \cdot 10^{-1}$	$2.02 \cdot 10^{-2}$	$1.43 \cdot 10^{-1}$	$2.02 \cdot 10^{-2}$
		CPU	1.08		$7.23 \cdot 10^{-3}$	
Assuming Clayton	ρ	1	$1.36 \cdot 10^{-1}$	$1.87 \cdot 10^{-2}$	$8.42 \cdot 10^{-2}$	$7.14 \cdot 10^{-3}$
		2	$9.98 \cdot 10^{-2}$	$1.00 \cdot 10^{-2}$	$1.01 \cdot 10^{-1}$	$1.03 \cdot 10^{-2}$
		3	$9.86 \cdot 10^{-2}$	$9.79 \cdot 10^{-3}$	$1.19 \cdot 10^{-1}$	$1.43 \cdot 10^{-2}$
		4	$1.12 \cdot 10^{-1}$	$1.26 \cdot 10^{-2}$	$1.49 \cdot 10^{-1}$	$2.22 \cdot 10^{-2}$
	ν	1	$2.44 \cdot 10^{-1}$	$6.02 \cdot 10^{-2}$	$3.52 \cdot 10^{-1}$	$1.24 \cdot 10^{-1}$
		2	$3.20 \cdot 10^{-1}$	$1.03 \cdot 10^{-1}$	$3.18 \cdot 10^{-1}$	$1.01 \cdot 10^{-1}$
		3	$3.23 \cdot 10^{-1}$	$1.05 \cdot 10^{-1}$	$2.80 \cdot 10^{-1}$	$7.88 \cdot 10^{-2}$
		4	$2.96 \cdot 10^{-1}$	$8.79 \cdot 10^{-2}$	$2.16 \cdot 10^{-1}$	$4.70 \cdot 10^{-2}$
		CPU	$5.65 \cdot 10^{-2}$		$1.25 \cdot 10^{-3}$	



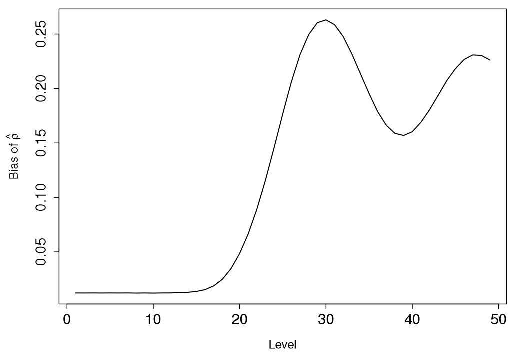
(a) Bias ratio of $\hat{\tau}$. (b) Bias ratio of $\hat{\lambda}$.

(c) Bias ratio of $\hat{\tau}$. (d) Bias ratio of $\hat{\lambda}_L$.

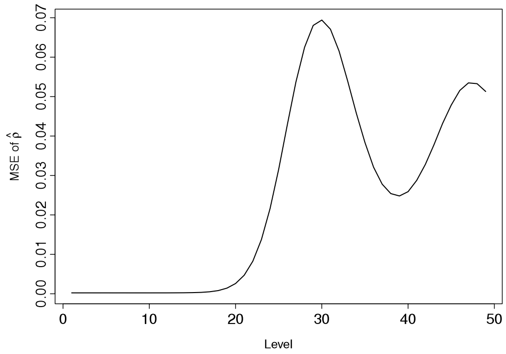
Fig. 5. Bias ratios of the SSP (numerator) versus the SP (denominator) estimates of the Kendall's τ (to the left) and (lower) tail dependence λ (λ_L) coefficients of Model 2(b) mixed with Model 2(a) (top row) and Model 2(c) mixed with Model 2(a) (bottom row) as functions of the mixing probability p, averaged over each of the four levels.

are well estimated, one may hope that the imprecise estimates for the top levels do not affect the lower order dependences too much.

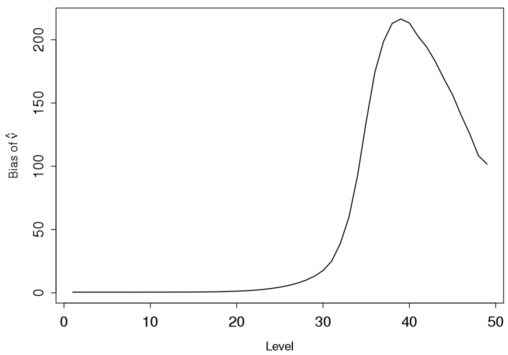
To study this, we have generated three samples of size $n = 5000$ from one of the estimated models, and compared certain lower order dependence characteristics of these samples to those of the true distribution, more specifically the (unconditional) pairwise Kendall's τ coefficients and the 90%, 95%, 97.5% and 99% quantiles of $u_i + u_j$, for all pairs (i, j).



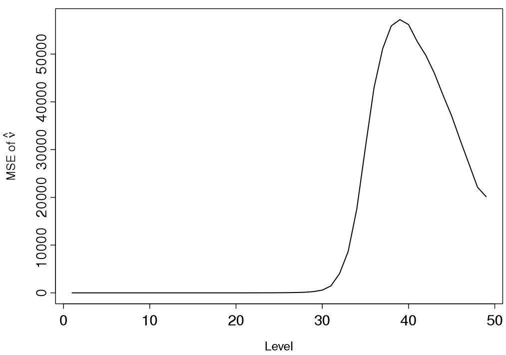
(a) Bias of $\hat{\rho}$.



(b) MSE of $\hat{\rho}$.



(c) Bias of $\hat{\nu}$.



(d) MSE of $\hat{\nu}$.

Fig. 6. Bias and MSE of the parameter estimates for the Student's t-vine in 50 dimensions, with low dependence ($\rho = 0.2$), averaged over each level.

As the true Kendall's τ coefficients and quantiles are difficult to compute analytically, we approximated them by the empirical equivalents based on a sample of size $n = 100{,}000$ from the true model. The results are displayed in Figs. 7 and 8, respectively. The values are averaged over the level the corresponding conditional dependence belongs to. For instance, the dependence between U_1 and U_{40} is modeled conditionally at level 39. The quantiles and Kendall's τ coefficients for this pair therefore contribute to the mean at level 39 in the plots. As expected, the Kendall's τ coefficients and quantiles corresponding to the lower levels of the structure appear to be close to the true values. However, they are fairly good also for the pairs modeled in the top levels, even though the conditional dependence between these pairs has been highly underestimated. Hence, the parameter estimates for the upper levels do not seem to have that much effect on the lower order dependences. This is an argument for truncating large structures after a certain level, letting the top level copulae be independence copulae, as proposed by Brechmann et al. [5].

Since the bias and MSE curves were rather different from what we had anticipated, we repeated the above experiments with a t-copula of dimension $d = 50$, all pairwise correlations set to 0.2 and $\nu = 6$. That is actually a special case of a D-vine with t-copulae. At a given level ℓ, the correlation parameters are the corresponding partial correlations, and the number of degrees of freedom is $\nu + \ell - 1 = 5 + \ell$. We simulated from this D-vine as described earlier, and estimated its parameters using the SSP estimator. Since the correlation parameters decrease with each level, whereas the degrees of freedom increase, it is more natural to consider the relative bias and MSE in this case, i.e. $Bias(\hat{\theta}_i, \theta_i)/\theta_i$ and $MSE(\hat{\theta}_i, \theta_i)/\theta_i^2$. We have also computed the average Kendall's τ coefficients and quantiles per level, based on simulations from estimated distributions. This is an unnecessarily cumbersome way of estimating the parameters of a t-copula. The purpose of this experiment was only to investigate whether either the simulation or the estimation routines are flawed on an example for which we know the true Kendall's τ coefficients and quantiles. Therefore, we have chosen not to show the actual results; we simply noted that they were reassuring (the interested reader may consult [16]). The relative bias and MSE increased steadily with the level of the structure, as one would expect. Moreover, all the pairwise Kendall's τ coefficients fluctuated around the true value of about 0.128, and likewise for the quantiles.

5. Conclusion

There are various estimators for the parameters of a pair-copula construction. The aim of this work has been to study and compare two of the most common ones, namely the semiparametric and stepwise semiparametric estimators. Both of these are based on transforming the original data with their empirical distribution functions. While the SP procedure consists

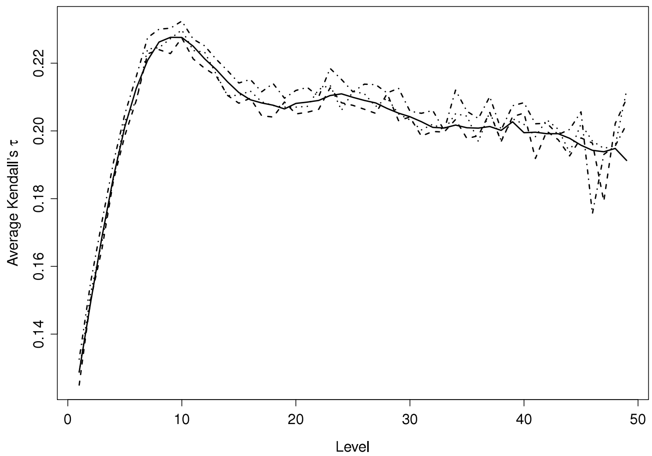


Fig. 7. Empirical pairwise Kendall's τ coefficients for 50,000 samples from the true distribution (connected line) and three samples generated from the estimated vine (dashed and dotted lines), averaged over the level the corresponding modeled conditional dependence belongs to.

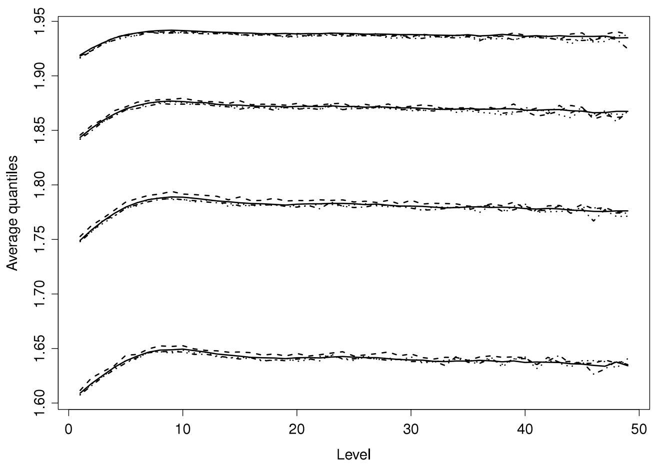


Fig. 8. Estimated 90%, 95%, 97.5% and 99% quantiles for 50,000 samples from the true distribution (connected lines) and three samples generated from the estimated vine (dashed and dotted lines), averaged over the level the corresponding modeled conditional dependence belongs to.

in estimating all parameters simultaneously, the SSP procedure addresses one level at a time, plugging in estimates from preceding levels.

In order to compare the two estimators, we have carried out a simulation study based on D-vines, a subset of PCCs. Except for one example, all the models considered are five-dimensional. Moreover, they are based on at least one out of three copula types, namely the Gaussian, the Student's t and the Clayton copulae. Varying the degree of dependence and the sample size, we have studied the effect on the two estimators.

Generally, the finite sample bias and MSE of the SSP estimator are higher than its competitor's, reflecting its lower asymptotic efficiency. The difference between the two estimators increases, in favor of SP, with the degree of dependence, and also with the level when the dependence is strong. Most likely, this is due to SSP's greater sensitivity toward the repeated transformations of the pseudo observations. Nonetheless, the performance of the SSP estimator is overall rather good compared to SP. Moreover, the former is consistently faster than the latter, especially when the number of parameters is high, as one would expect.

When the sample size decreases, both estimators' variances increase. Based on 500 observations, the parameter estimates for the five-dimensional vines are still rather precise. A sample size of $n = 50$ is insufficient, especially for the degrees of freedom parameter of Student's t-copulae. Both estimators struggle with that parameter, SSP in particular. For the remaining parameters, the difference between the estimators becomes smaller with the sample size. That is also the case when the model is not correctly specified. Neither the SP nor the SSP estimators are particularly robust toward misspecification of the pair-copulae constituting the PCC, but in most cases, the former appears to suffer more than the latter. Hence, the gain from using the more time consuming SP estimator may be reduced for small samples and inaccurate model assumptions.

For high-dimensional problems, the SP estimator is simply too expensive, and would probably be numerically unstable, whereas SSP estimation still is doable. Used on a 50-dimensional Student's t-vine, the resulting estimates are quite good up to level 20–30. After that, the finite sample bias and MSE explode. Estimation of such high order dependences is unfortunately numerically highly challenging. However, this does not seem to affect the corresponding lower order dependences. Despite the erroneous estimates for the top levels, the estimated distribution is in fact rather similar to the true one. This is an incentive to truncate large structures after a certain level, letting the copulae of the top levels be the independence copula.

Overall, the SSP estimator is well suited to set starting values for the SP one, in small to medium sized problems. Furthermore, the extra time spent on SP estimation is not necessarily worthwhile, especially when the dependence is not too strong, the sample size is low, or the model is misspecified. In cases of very strong or extreme dependence, a subsequent SP estimation is necessary, whereas small-sized samples require other types of estimators. When the number of parameters becomes large, SSP estimation is the only alternative in practice. Of course, we have not considered cases of extreme dependence, for which the SSP estimator is likely to be outperformed. Also, we have restricted our attention to D-vines. The relative behavior of the two estimators for C-vines and other regular vines may be a subject for future work, although we believe the results will be similar to the ones for D-vines. All in all, this study supports the use of SSP in most applications.

Acknowledgments

This work is funded by Statistics for Innovation, $(\text{sfi})^2$. I thank my supervisors Arnoldo Frigessi and Kjersti Aas for the helpful discussions and comments. I also thank the referees and the Associate Editor for their help to improve this paper with their good comments and suggestions.

References

[1] K. Aas, C. Czado, A. Frigessi, H. Bakken, Pair-copula constructions of multiple dependence, Insurance: Mathematics and Economics 44 (2) (2009).
[2] T. Bedford, R. Cooke, Probabilistic density decomposition for conditionally dependent random variables modeled by vines, Annals of mathematics and Artificial Intelligence 32 (2001) 245–268.
[3] T. Bedford, R. Cooke, Vines — a new graphical model for dependent random variables, Annals of Statistics 30 (4) (2002) 1031–1068.
[4] D. Berg, Copula goodness-of-fit testing: an overview and power comparison, European Journal of Finance 15 (2009) 675–701.
[5] E.C. Brechmann, C. Czado, K. Aas, Truncated regular vines in high dimensions with application to financial data, 2010 (submitted for publication).
[6] D. Clayton, A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, Biometrika 65 (1978) 141–151.
[7] C. Czado, A. Min, Bayesian inference for multivariate copulas using pair-copula constructions, Journal of Financial Econometrics 8 (2010) 511–546.
[8] P. Embrechts, A.J. McNeil, D. Straumann, Correlation: pitfalls and alternatives, Risk 12 (1999) 69–71.
[9] E. Frees, E. Valdez, Understanding relationships using copulas, North American Actuarial Journal 2 (1998) 1–25.
[10] C. Genest, Frank's family of bivariate distributions, Biometrika 74 (1987) 549–555.
[11] C. Genest, M. Gendron, M. Bourdeau-Brien, The advent of copulas in finance, European Journal of Finance 15 (2009) 609–618.
[12] C. Genest, K. Ghoudi, L. Rivest, A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, Biometrika 82 (1995) 543–552.
[13] C. Genest, B. Rémillard, D. Beaudoin, Goodness-of-fit tests for copulas: a review and power study, Insurance: Mathematics and Economics 44 (2009) 199–213.
[14] C. Genest, L.-P. Rivest, Statistical inference procedures for bivariate archimedean copulas, Journal of the American Statistical Association 88 (1993) 1034–1043.
[15] I. Hobæk Haff, Parameter estimation for pair-copula constructions, Technical Report SAMBA/36/10, Norsk Regensentral, 2010.
[16] I. Hobæk Haff, Comparing estimators for pair-copula constructions, Technical Report SAMBA/22/11, Norsk Regnesentral, 2011.
[17] I. Hobæk Haff, K. Aas, A. Frigessi, On the simplified pair-copula construction — simply useful or too simplistic? Journal of Multivariate Analysis 101 (2010) 1296–1310.
[18] H. Joe, Chapter families of m-variate distributions with given margins and $m(m-1)/2$ dependence parameters, in: Distributions with Fixed Marginals and Related Topics, IMS, Hayward, CA, 1996.
[19] H. Joe, Multivariate Models and Dependence Concepts, Chapman & Hall, London, 1997.
[20] H. Joe, Asymptotic efficiency of the two-stage estimation method for copula-based models, Journal of Multivariate Analysis 94 (2005) 401–419.
[21] H. Joe, H. Li, A.K. Nikoloulopoulos, Tail dependence functions and vine copulas, Journal of Multivariate Analysis 101 (2010) 252–270.
[22] H. Joe, J. Xu, The estimation method of inference functions for margins for multivariate models, Technical Report 166, University of British Columbia, Department of Statistics, 1996.
[23] G. Kim, M. Silvapulle, P. Silvapulle, Comparison of semiparametric and parametric models for estimating copulas, Computational Statistics and Data Analysis 51 (2007) 2836–2850.
[24] O. Kolbjørnsen, M. Stien, D-vine creation of non-gaussian random field, in: Procedings of the Eight International Geostatistics Congress, GECAMIN Ltd., 2008, pp. 399–408.
[25] D. Kurowicka, R. Cooke, Uncertainty Analysis with High Dimensional Dependence Modelling, Wiley, New York, 2006.
[26] R. Nelsen, An Introduction to Copulas, in: Lecture notes in Statistics, vol. 139, Springer Verlag, New York, 1999.
[27] D. Oakes, A model for association in bivariate survival data, Journal of the Royal Statistical Society, Series B 44 (1982) 414–422.
[28] J. Shih, T. Louis, Inferences on the association parameter in copula models for survival data, Biometrics 51 (1995) 1384–1399.
[29] A. Sklar, Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de statistique de l'Université de Paris, 8, 1959.
[30] H. Tsukahara, Semiparametric estimation of copula models, The Canadian Journal of Statistics 33 (2005) 357–375.